Studies on Tenuring Collection Times for a Generational Garbage Collector∗

Xufeng Zhao1,2 Cunhua Qian2 Syouji Nakamura3

1Department of Business Administration
Aichi Institute of Technology, Toyota 470-0392, Japan
2School of Economics and Management
Nanjing University of Technology, Nanjing 210009, China
3Department of Human Life and Information
Kinjo Gakuin University, Nagoya 463-8521, Japan

Abstract It is an important problem to analyze the collection costs and determine tenuring collection times to meet the pause time goal for a generational garbage collector. From such viewpoints, this paper analyzes the costs suffered for collections according to the working schemes firstly, and then, proposes one garbage collection model. Garbage collections occur at a nonhomogeneous Poisson process, and tenuring collections are made at time $T$ and at level $K$. Applying the techniques of degradation processes and continuous wear processes, the expected cost rate is derived, four cases of optimal policies for the model are discussed analytically and numerically.

Keywords Garbage Collection; Tenuring Collection; Minor Collection; Continuous Process; Optimal Policy

1 Introduction

In recent years, generational garbage collection [1-5] has been popular with programmers for the reason that it can be made more efficient. Compared with classical tracing collectors, e.g., reference counting collector, mark-sweep collector, mark-compact collector and copying collector, a generational garbage collector is effective in applications of the computer programs with character that it is unnecessary to mark or copy all active data of the whole heap for every collection, i.e., the collector concentrates effort on those objects most likely to be garbage. Based on the weak generational hypothesis [2] which asserts that most objects are short-lived after their allocation, a generational garbage collector segregates objects by age into several regions called generations or multiple generations. The survival rates of younger generations are always much lower than those of older ones, which means that younger generations are more likely to be garbage and can be collected more frequently than older ones. Although such collections cost much shorter time than those of full collections, the problems of pointers from older generations to younger ones and the size of root sets for younger generations will become more

∗This project is supported by Grant-in-Aid for Scientific Research (C) of Japan Society for the Promotion of Science under Grant No. 22500897; National Natural Science Foundation of China (70471017, 70801036); Humanities and Social Science Research Foundation of China (05JA630027).
complicated. For these reasons, many collectors are limited to just two or three genera-
tions [1, p.153]. For instance, the collector which is used in Sun’s HotSpot Java Virtual
Machine (JVM) manages heap space for both young and old generations [5]: new objects
space Eden, two equal survivor spaces SS♯1 and SS♯2 for surviving objects, and tenured
objects space Old (Tenured), where Eden, SS♯1 and SS♯2 are for young generation, and
Old (Tenured) is for old one.

A generational garbage collector uses minor collection and tenuring collection for
young generation and major collection for multi-generations. However, for every garbage
collection, the manner of stop and copy pauses all application threads to collect the
garbage. The duration of time in which a garbage collector has worked is called pause
time [1, p.148], which is an important parameter for interactive systems, and depends
largely upon the volume of surviving objects and the type of collections. That is, pause
time suffered for minor collections increases with the number of collections and is less
than that of tenuring collections, major collection pause time is the longest among the
three ones.

There have been very few research papers that studied analytically optimal collection
times for a generational garbage collector. Most problems were concerned with several
ways to introduce garbage collection methods and how to tune the garbage collector,
which are complex to operate due to random accesses of programs in the memory [6-10].
We propose that garbage collection is a stochastic decision making process and should be
analyzed by the theory of stochastic processes. As the applications of damage models, a
garbage collection model for a database in the computer system [11] was studied, but the
theoretical point of collection was not considered essentially, and optimal policies for a
generational garbage collector with tenuring threshold [12] according to practical working
schemes were studied recently, however, it is may be difficult to inspect the survivor rates
exactly at collection times.

Furthermore, increase in objects might be unclear at discrete times for the high fre-
cquency of computer processes, it would be more practical to assume that surviving objects
that should be copied increase with time continuously and roughly according to some
mathematical laws. On the other hand, degradation processes [13] and continuous wear
processes [14] play an important role in describing the degradation of the unit. Life distrib-
ution properties, failure rates and some maintenance policies subjected to such processes
were summarized in [13, 14] and such processes has been applied to various fields such
as material, physics, finance and statistics [15-18].

This paper considers a pause time goal which is called time cost or cost for simplicity.
Our problem is to obtain optimal tenuring collection times which minimize the expected
cost rates. First, the working schemes and some analyses about collection costs are given.
Second, garbage collections occur at a nonhomogeneous Poisson process, and tenuring
collections are made at time $T$ and at level $K$. Applying the techniques of degradation
processes and continuous wear processes, the expected cost rate is derived, four cases of
optimal policies for the model are discussed analytically. Fourth, numerical examples of
the policies are given.
2 Models and Optimal Policies

2.1 Working Schemes

Detailed working schemes of a generational garbage collector in general which has been introduced in [1, 5, 12] are given as following steps:

1. New objects are allocated in Eden.
2. When the first minor collection occurs, surviving objects are copied from Eden into \( SS_1 \).
3. When the second minor collection occurs, surviving objects from Eden and \( SS_1 \) are copied into \( SS_2 \).
4. In the fashions of 1-3, minor collection copies surviving objects between the two survivor spaces until they become tenured, i.e., tenuring collection occurs when some parameter meets the tenuring threshold, and then, the older or oldest objects are copied into Old.
5. When Old fills up, major collection of the whole heap occurs, and surviving objects from Old are kept in Old, while objects from Eden and survivor space are kept in a survivor space.

Based on [2], from the viewpoints of collection efficiency, new objects can be tenured only if they survive at least one minor collection, because objects that survive two minor collections are much less than those survive just one minor collection. In other words, surviving objects are likely to reduce slightly with the number of minor collections beyond two. Although major collection spends much time, Old will be filled with tenured objects slowly, especially when the tenuring threshold is high and the survivor rate is low, that is, major collection occurs rarely in this case. So that, this paper concentrates only on minor and tenuring collections and considers tenuring collections as renewal points of the collection processes, that is, after tenuring collection, the same collection cycle begins with step 1: \( 1 \rightarrow 2 \rightarrow 3 \rightarrow 4 \rightarrow 1 \rightarrow \cdots \).

It is assumed that the total volume of surviving objects in Eden and survivor space at time \( t \) is \( Z(t) = A(t) + \sigma B(t) \) with distribution \( \Pr\{Z(t) \leq x\} = W(t,x) \), where \( A(t) \) and \( B(t) \) are random variables. Letting \( c_S + c_M(x) \) be the cost suffered for every minor collection, where \( c_S \) is the constant cost of scanning surviving objects and \( x \) is the surviving objects that should be copied, \( c_M(x) \) increases with \( x \) and \( c_M(0) \equiv 0 \). \( c_K, c_T \) and \( c_N \) (\( c_K, c_T, c_N > c_S + c_M(K) \)) be the cost suffered for tenuring collection when surviving objects have exceeded a threshold level \( K \) (\( 0 < K < \infty \)), at time \( T \) and at \( N \)th collection, respectively. Then, the expected cost of minor collection at time \( t \) is

\[
C(t,K) = \frac{1}{W(t,K)} \int_0^K \left[ c_S + c_M(x) \right] dW(t,x),
\]

where \( C(0,K) \equiv 0 \). It is clear that if \( -W'(t,x)/W(t,x) \) increases with \( t \) for any \( x \geq 0 \), \( C(t,K) \) increases with \( t \) for any \( K \geq 0 \).

2.2 Modeling and Optimization

It is assumed that garbage collections occur at a nonhomogeneous Poisson process with an intensity function \( \lambda(t) \) and a mean-value function \( R(t) \equiv \int_0^t \lambda(u)du \). Then, the
probability that collections occur exactly \( j \) times in \((0,t]\) is
\[
H_j(t) ≡ \frac{[R(t)]^j}{j!} e^{-R(t)} \quad (j = 0, 1, 2, \cdots).
\]

Letting \( F_j(t) \) denote the probability that collections occur at least \( j \) times in \((0,t]\). Then,
\[
F_j(t) = \int_0^t H_{j-1}(u)\lambda(u)du = \sum_{i=j}^{\infty} H_i(T), \quad (j = 1, 2, \cdots), \tag{2}
\]

where \( F_0(t) ≡ 1 \).

Suppose that minor collections are made when the garbage collector begins to work, tenuring collection is made at a planned time \( T \) \((0 < T \leq \infty)\), or when surviving objects have exceeded a threshold level \( K \) \((0 < K \leq \infty)\), whichever occurs first. Then, the mean time to tenuring collection is
\[
E(L) = TW(T,K) + \int_0^T t\mathcal{W}(t,K)dt = \int_0^T W(t,K)dt, \tag{3}
\]

where \( \mathcal{W}(t,x) ≡ 1 - V(t,x) \) for any distribution \( V(t,x) \).

The expected cost suffered for minor collections until tenuring collection is
\[
C_M(T,K) = W(T,K)\sum_{j=1}^{\infty} \int_0^T C(t,K)dF_j(t) + \int_0^T \left[ \sum_{j=1}^{\infty} \int_0^T C(u,K)dF_j(u) \right]d\mathcal{W}(t,K)
= \sum_{j=1}^{\infty} \int_0^T C(t,K)W(t,K)dF_j(t) = \int_0^T \lambda(t)C(t,K)W(t,K)dt. \tag{4}
\]

Then, the expected cost until tenuring collection is
\[
E(C) = c_K - (c_K - c_T)W(T,K) + \int_0^T \lambda(t)C(t,K)W(t,K)dt. \tag{5}
\]

Therefore, from (3) and (5), by using the renewal reward process [19], the expected cost rate is
\[
C(T,K) = \frac{c_K - (c_K - c_T)W(T,K) + \int_0^T \lambda(t)C(t,K)W(t,K)dt}{\int_0^T W(t,K)dt}. \tag{6}
\]

It can be seen that \( C(T,K) \) includes the following collection polices:

- Tenuring collection is made at time \( T \) for a given \( K \), the reason why making such a policy is \( c_T < c_K \).
- Tenuring collection is made at level \( K \) for a given \( T \). In this case, \( c_K < c_T \).
- Tenuring collection is made only at time \( T \) or only at level \( K \). In these two cases, \( c_K = c_T \).
2.2.1 Optimal $T^*$

When $c_T < c_K$, we find an optimal $T^*$ which minimizes $C(T,K)$ in (6) for a given $K$. Letting $r(t,x)$ be the failure rate of $W(t,x)$, i.e., $r(t,x) ≡ -W'(t,x)/W(t,x)$ [20]. Then, differentiating $C(T,K)$ with respect to $T$ and setting it equal to zero,

\[
(c_K - c_T) \left[ r(T,K) \int_0^T W(t,K)dt - \bar{W}(T,K) \right] \\
+ \int_0^T \left[ \lambda(T)C(T,K) - \lambda(t)C(t,K) \right] W(t,K)dt = c_T.
\]

Letting $L_1(T,K)$ be the left-hand side of (7),

\[ L_1(0,K) \equiv \lim_{T \to 0} L_1(T,K) = 0, \]

\[ L_1'(T,K) = (c_K - c_T) r'(T,K) \int_0^T W(t,K)dt \\
+ \left[ \lambda'(T)C(T,K) + \lambda(T)C'(T,K) \right] \int_0^T W(t,K)dt. \]

Thus, if $r(T,K)$ and $\lambda(t)$ increase with $t$, then the left-hand side of (7) increases with $t$ from 0. Therefore, there exists a unique optimal $T^*$ ($0 < T^* \leq \infty$) which satisfies (7), and the resulting cost rate is

\[ C(T^*,K) = (c_K - c_T) r(T^*,K) + \lambda(T^*)C(T^*,K). \]

2.2.2 Optimal $K^*$

When $c_K < c_T$, we find an optimal $T^*$ which minimizes $C(T,K)$ in (6) for a given $T$. Letting $w(t,x)$ be a density function of $W(t,x)$, i.e., $w(t,x) \equiv dW(t,x)/dx$. Then, differentiating $C(T,K)$ with respect to $K$ and setting it equal to zero,

\[
(c_T - c_K) \left[ \lambda_1(T,K) \int_0^T w(t,K)dt - \bar{w}(T,K) \right] \\
+ \int_0^T \left[ \lambda_2(T,K) - \lambda(t)C(t,K) \right] W(t,K)dt = c_K,
\]

where

\[ \lambda_1(T,K) \equiv \frac{w(T,K)}{\int_0^T w(t,K)dt}, \quad \lambda_2(T,K) \equiv \frac{[c_S + c_M(K)] \int_0^T \lambda(t)w(t,K)dt}{\int_0^T w(t,K)dt}. \]

Letting $L_2(T,K)$ be the left-hand side of (8),

\[ L_2(0,K) \equiv \lim_{K \to 0} L_2(T,K) = 0, \]

\[ L_2'(T,K) = (c_T - c_K) \lambda_1(T,K) \int_0^T w(t,K)dt \\
+ \lambda_2(T,K) \int_0^T W(t,K)dt. \]
Thus, if $Q_1(T, K)$ and $Q_2(T, K)$ increase with $K$, then the left-hand side of (8) increases with $K$ from 0. Therefore, there exists a unique optimal $K^*$ ($0 < K^* \leq \infty$) which satisfies (8), and the resulting cost rate is

$$C(T, K^*) = (c_T - c_K)Q_1(T, K^*) + Q_2(T, K^*).$$

### 2.2.3 Optimal $\bar{T}^*$

When $c_K = c_T$, putting that $K = \infty$ in (6), the expected cost rate is

$$\bar{C}(T) \equiv \lim_{K \to \infty} C(T, K) = \frac{1}{T} \left[ \int_0^T \lambda(t)C(t, \infty)dt + c_T \right], \quad (9)$$

From (7), if $\lambda(t)$ increases with $t$, then an optimal tenuring collection time $\bar{T}^*$ which minimizes (9) is given by a unique solution of the equation

$$\int_0^T \left[ \lambda(t)C(T, \infty) - \lambda(t)C(t, \infty) \right] dt = c_T, \quad (10)$$

and the resulting cost rate is

$$\bar{C}(\bar{T}^*) = \lambda(\bar{T}^*)C(\bar{T}^*, \infty).$$

In particular, when $\lambda(t) = \lambda$, (10) becomes

$$\int_0^\infty \left[ \int_0^T [W(t, x) - W(T, x)] \right] dcM(x) = \frac{c_T}{\lambda}, \quad (11)$$

which increases with $T$.

### 2.2.4 Optimal $\bar{K}^*$

When $c_K = c_T$, putting that $T = \infty$ in (6), the expected cost rate is

$$\bar{C}(K) = \lim_{T \to \infty} C(T, K) = \frac{\int_0^\infty \lambda(t)C(t, K)W(t, K)dt + c_K}{\int_0^\infty W(t, K)dt}, \quad (12)$$

From (8), if $Q_2(\infty, K)$ increases with $K$, then an optimal tenuring collection time $\bar{K}^*$ which minimizes (12) is given by a unique solution of the equation

$$\int_0^\infty \left[ Q_2(\infty, K) - \lambda(t)C(t, K) \right] W(t, K)dt = c_K, \quad (13)$$

and the resulting cost rate is

$$\bar{C}(\bar{K}^*) = Q_2(\infty, \bar{K}^*).$$

In particular, when $\lambda(t) = \lambda$, (13) becomes

$$\int_0^\infty \left[ \int_0^K [W(t, x)dcM(x) \right] dt = \frac{c_K}{\lambda}, \quad (14)$$

which increases with $K$. 

---

Studies on Tenuring Collection Times for a Generational Garbage Collector 289
Table 1: Optimal $T^*$ and $C(T^*, K)/c_T$ when $c_S/c_T = 0.1$ and $\lambda = \mu = \sigma = 1$.

<table>
<thead>
<tr>
<th>$K$</th>
<th>$c_K/c_T$</th>
<th>$c_M/c_T = 0.01$</th>
<th>$c_M/c_T = 0.05$</th>
<th>$c_M/c_T = 0.1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>3.73</td>
<td>0.4723</td>
<td>4.09</td>
<td>3.57</td>
</tr>
<tr>
<td>3</td>
<td>3.25</td>
<td>0.5473</td>
<td>3.06</td>
<td>2.87</td>
</tr>
<tr>
<td>5</td>
<td>2.80</td>
<td>0.5891</td>
<td>2.69</td>
<td>2.56</td>
</tr>
<tr>
<td>10</td>
<td>5.49</td>
<td>0.4851</td>
<td>5.32</td>
<td>0.4722</td>
</tr>
</tbody>
</table>

Table 2: Optimal $K^*$ and $C(T, K^*)/c_K$ when $c_S/c_K = 0.1$ and $\lambda = \mu = \sigma = 1$.

<table>
<thead>
<tr>
<th>$T$</th>
<th>$c_T/c_K$</th>
<th>$c_R/c_K = 0.01$</th>
<th>$c_R/c_K = 0.05$</th>
<th>$c_R/c_K = 0.1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>2.69</td>
<td>0.2990</td>
<td>5.77</td>
<td>4.50</td>
</tr>
<tr>
<td>3</td>
<td>6.37</td>
<td>0.3160</td>
<td>5.40</td>
<td>4.42</td>
</tr>
<tr>
<td>5</td>
<td>5.63</td>
<td>0.3252</td>
<td>5.18</td>
<td>4.36</td>
</tr>
</tbody>
</table>

3 Numerical Examples

We compute numerical examples of the models discussed above for $Z(t) = \mu + \sigma B(t)$ when $B(t)$ is normally distributed with mean 0 and variance $t$ or for $Z(t) = A(t)t$ when $A(t)$ is normally distributed with mean $\mu$ and variance $\sigma^2/1$, that is,

$$W(t, x) = \Phi \left( \frac{x - \mu t}{\sigma \sqrt{t}} \right), \quad (15)$$

where $\Phi(x)$ is the standard normal distribution with mean 0 and variance 1, i.e., $\Phi(x) \equiv (1/\sqrt{2\pi}) \int_{-\infty}^{x} e^{-u^2/2} du$.

Table 3: Optimal $\tilde{T}^*$, $\tilde{C}(\tilde{T}^*)/c_T$, $\tilde{K}^*$ and $\tilde{C}(\tilde{K}^*)/c_K$ when $\lambda = \mu = \sigma = 1$.

<table>
<thead>
<tr>
<th>$c_M/c_T = 0.1$</th>
<th>$c_M/c_T = 0.1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\tilde{T}^*$</td>
<td>$\tilde{C}(\tilde{T}^*)/c_T$</td>
</tr>
<tr>
<td>0.01</td>
<td>14.24</td>
</tr>
<tr>
<td>0.02</td>
<td>10.11</td>
</tr>
<tr>
<td>0.03</td>
<td>8.28</td>
</tr>
<tr>
<td>0.04</td>
<td>7.20</td>
</tr>
<tr>
<td>0.05</td>
<td>6.46</td>
</tr>
<tr>
<td>0.06</td>
<td>5.91</td>
</tr>
<tr>
<td>0.07</td>
<td>5.49</td>
</tr>
<tr>
<td>0.08</td>
<td>5.15</td>
</tr>
<tr>
<td>0.09</td>
<td>4.86</td>
</tr>
<tr>
<td>0.10</td>
<td>4.62</td>
</tr>
</tbody>
</table>
References